George G. Vega Yon, Ph.D.

Data Science, Statistical Computing, & Complex Systems Modeling

I'm an accomplished **data scientist** with over **10 years of experience** with multiple software packages and scientific publications. My software has been downloaded over **half a million times**, and my papers have over **200 citations**. My work has spanned data science, economics, network science, statistics, and phylogenetics. I seek new challenges, enjoy collaborating with others, and excel at mentoring.

Mobile	+1 (six two six) 381 8171
e-mail	g.vegayon@gmail.com
website	ggvy.cl
Code	github.com/gvegayon
Linkedin	www.linkedin.com/in/georgevegayon/

I am a successful scientific software developer with over **half a million downloads**

Education

Ph.D. in Biostatistics (concentration in Stat. Comp.) University of Southern California (2020). Dissertation *"Essays on Bioinformatics and Social Network Analysis: Statistical and Computational Methods for Complex Systems"*

M.Sc. in Social Sciences (Economics) California Institute of Technology (2016)

Master in Economics and Public Policy, Universidad Adolfo Ibáñez (2011)

BS. in Social Sciences and **BS. in Business Administration**, Universidad Adolfo Ibáñez (2011)

Professional Experience

University of Utah, November 2021–Present Division of Epidemiology *Research Assistant Professor*. Division of Population Health Sciences, *Adjunct Assistant Professor*.

University of Southern California, 2015–November 2021 Department of Preventive Medicine *Research Programmer*. As a senior research staff member, my responsibilities include: Provide technical support on statistical computing, e.g., HPC, run training sessions on scientific software development, and writing scientific papers. Since August 2020, I co-instruct the Department's Introduction to Data Science class.

Chilean Pension Supervisor, August 2011– August 2014 Research Division *Analyst*. Statistical and econometric analysis on the Chilean unemployment insurance, statistical software development, serving as a bridge between the IT and Research divisions.

Universidad Adolfo Ibáñez, January 2011–June 2012. School of Government *Adjunct Professor*. Taught Introductory courses of Economics, Microeconomics and Statistical computing with Stata.

Software (selected)

[1] **George G. Vega Yon**. *aphylo: Statistical Inference of Annotated Phylogenetic Trees* (2022). R package version 0.2-1 URL: https://cran.r-project.org/package=aphylo.
`downloads 6058`

[2] **George G. Vega Yon**. *rgexf: Build, Import and Export GEXF Graph Files* (2020). R package version 0.16.0. URL: https://CRAN.R-project.org/package=rgexf. `downloads 601K`

[3] **George G. Vega Yon**, Thomas Valente. *netdiffuseR: Analysis of Diffusion and Contagion Processes on Networks* (2020). R package version 1.22.0. URL: https://github.com/USCCANA/netdiffuseR. `downloads 38K`
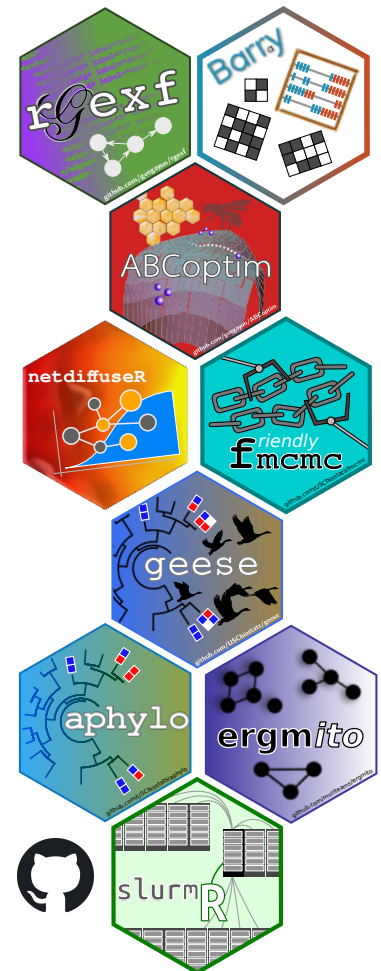
[4] **George G. Vega Yon**, Kayla de la Haye. *ergmito: Exponential Random Graph Models for Small Networks* (2020). R package version 0.3-0. URL: `https://cran.r-project.org/package=ergmito`. downloads `19K`

[5] **George G. Vega Yon**. *slurmR: A Lightweight Wrapper for 'Slurm'* (2020). R package version 0.4-1. URL: `https://CRAN.R-project.org/package=slurmR`. downloads `25K`

[6] **George G. Vega Yon**. *fmcmc: A friendly MCMC framework* (2020). R package version 0.3-0. URL: `https://CRAN.R-project.org/package=fmcmc`. downloads `21K`

Academic Publications (selected)

[1] **George G. Vega Yon**. "Power and Multicollinearity in Small Networks: A Discussion of "Tale of Two Datasets: Representativeness and Generalisability of Inference for Samples of Networks" by Krivitsky, Coletti & Hens". In: *Journal of The American Statistical Association* (2023). toappear.

[2] **George G. Vega Yon**, Mary Jo Pugh, and Thomas W. Valente. *Discrete Exponential-Family Models for Multivariate Binary Outcomes*. Nov. 2022. arXiv: `2211.00627 [cs, stat]`. (Visited on 11/02/2022).

[3] **George G. Vega Yon**, Andrew Slaughter, and Kayla de la Haye. "Exponential random graph models for little networks". In: *Social Networks* 64 (2021), pp. 225–238. URL: `https://doi.org/10.1016/j.socnet.2020.07.005`.

[4] **George G.. Vega Yon**, Duncan C. Thomas, John Morrison, Huaiyu Mi, et al. "Bayesian parameter estimation for automatic annotation of gene functions using observational data and phylogenetic trees". In: *PLOS Computational Biology* 17.2 (Feb. 2021), pp. 1–35. URL: `https://doi.org/10.1371/journal.pcbi.1007948`.

[5] **George G. Vega Yon** and Paul Marjoram. "fmcmc: A friendly MCMC framework". In: *Journal of Open Source Software* 4.39 (July 2019), p. 1427. URL: `http://joss.theoj.org/papers/10.21105/joss.01427`.

[6] **George G. Vega Yon** and Brian Quistorff. "parallel: A command for parallel computing". In: *The Stata Journal: Promoting communications on statistics and Stata* 19.3 (Sept. 2019), pp. 667–684. URL: `http://journals.sagepub.com/doi/10.1177/1536867X19874242`.

Honors and Professional Achievements

Technologies R, C++, LaTeX, SQL, Python, XML, regex, Stata, AWS, Gephi, Pajek, Mathematica, Git, Docker, Visual Studio Code, tensorflow, continuous integration, Slurm, Unix.

Awards Best paper, 72 ICA conference, 2022; Travel Grant, Society of Young Network Scientist, 2019; Fellowship, California Institute of Technology, 2014; Scholarship, Adolfo Ibáñes University, 2006.

Manuscript reviewer JASA, BMC Infectious Diseases, The Official Journal of The Society for Computational Economics, The R Journal, Social Networks, Journal of Mathematical Sociology, JOSS, Bioinformatics, Computer Methods and Programs in Biomedicine Update, SUNBELT Conference (2016), IC2S2 (2019–2021).

Misc Founder of the R Users Group in Chile (2013), co-organizer of the East LA R User Group (LAERUG).